

A new approach for sentiment and geopolitical stance analysis of Reddit comments on the Israel-Palestine conflict

Gopal D. Upadhye¹, Hrushikesh Joshi¹, Satpalsing Devising Rajput², Ranjana S. Jadhav³, Pranav Moralarwar¹, Tejas Muttalwar¹, Yash Mehendale¹, Siddharth Mane¹, Shivam Mishal¹

¹Department of Artificial Intelligence of Data Science, Vishwakarma Institute of Technology, Pune, India

²Department of Computer Science and Engineering, Pimpri Chinchwad University, Pune, India

³Department of Information Technology, Vishwakarma Institute of Technology, Pune, India

Article Info

Article history:

Received Jul 30, 2025

Revised Oct 8, 2025

Accepted Dec 6, 2025

Keywords:

Israel-Palestine conflict
geopolitical stance
Latent Dirichlet allocation
Natural language processing
valence aware dictionary and
sentiment reasoner
Public opinion analysis
Reddit
Sentiment analysis

ABSTRACT

The paper focuses on the persistent conflict between Israel and Palestine. of world discussion, with such sites as Reddit containing a great deal medium for public expression. This study uses natural language processing (NLP) including valence aware dictionary and sentiment reasoner (VADER) for sentimental analysis and latent Dirichlet allocation (LDA) for topic modelling of geopolitical positions in comments of 400,000 posts belonging to the topics subreddits. A sentiment analysis using VADER shows that negative sentiments observed and indicate that such sentiments prevail throughout fear of aggression, human suffering, and threat. media biases. LDA analysis for topic modelling. The process of LDA reveals such important topics as humanitarian ones. media discourses, polemizing and appeals for calm. Furthermore, the study which divides comments into geopolitical attitudes – support for or yet, Israelis, supporters of Palestine, opponents of both, and neutrals, offering a contingency comprehensive conception of public perspectives. The results emphasize the effectiveness of large-scale. Of the big data pre-processing techniques, NLP can be used to understand geopolitics and its intricate dynamics. educate policy makers, researchers, and media planners.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Gopal D. Upadhye

Department of Artificial Intelligence of Data Science, Vishwakarma Institute of Technology

Indiranagar, Bibwewadi, Pune, Maharashtra, 411037, India

Email: gopal.upadhye@vit.edu

1. INTRODUCTION

Israel/Palestine remains one of the most intractable political issues and some of the oldest roots are explored here. The existing conflict between the Israelis and Palestinians have spurred an extensive discursive practice on social networks particularly on Reddit [1]. As the amount of content being shared increases on the web, the relevance of understanding people's mood and topics being discussed in those posts rises as well [2], [3]. Using sentiment analysis, topic modeling, and geopolitical stance detection, the research aims at identifying the prevailing emotions, topics and international political leaning. The first research question of this study is also centered on evaluating the sentiment present in comments of Reddit. Sentiment analysis enables one to determine if the comments they contain are positive, negative or rather neutral [4], [5], thereby giving insight on how different users feel about the ongoing conflict. Furthermore, in the comments section the study identifies what topics of discussion these comments contain through topic modeling which relevant issues like humanitarian issues, media coverage, military tactics, and appeals for peace [6]. Last, according to the position of comments in the geopolitical conflict—either pro-Israel or pro-

Palestine, against both parties, or having no preference, the study provides a better perception of the comments' diversity present in such online discussions [7], [8]. But this study also has following difficulties. Adaptive work environment is challenging in part due to the fact that comments on social networks are often filled with sarcasm and irony or other forms of statements that are to certain extent ambiguous [4]. Moreover, the fact that the text of discussion in different subreddits could be entirely different, reflecting the peculiarities of the corresponding communities, while acknowledging this or that or a number of topics, could introduce biases into sentiment and/or stance analysis [9], [10]. However, this research aspires to contribute valuable insight with regard the emotion and political aspects involved in the Reddit discourse on the Israel-Palestine issue [11]. Compared to previous research, this study makes four contributions which are as follows: integrated framework- unlike prior studies that focused on sentiment or topics, this work combines sentiment analysis, topic modeling, and stance classification into a unified pipeline. Coherence-optimized topic modeling- instead of arbitrarily choosing topic numbers, the study validates latent Dirichlet allocation (LDA) results using coherence and perplexity scores, improving reliability. Explicit Stance Logic- Stances are divided into four clear categories (pro-Israel, pro-Palestine, opposition to both, and neutral), with heuristic+supervised methods ensuring reproducibility. Dataset Scale-Over 400,000 Reddit comments make this one of the largest datasets used in conflict-related natural language processing (NLP) research.

2. LITERATURE REVIEW

Hutto and Gilbert [1] introduced valence aware dictionary and sentiment reasoner (VADER), a lexicon model that efficiently analyzes sentiment in social media contexts. Al-Sarraj and Lubbad (2018) [8] explored the concept of sentiment analysis in identifying bias within discussion, utilizing classifiers such as support vector machine (SVM), while Liyih *et al.* [12] achieved strong results with a CNN-BiLSTM, though sarcasm remained difficult. K *et al.* (2023) [13] used machine learning applications to analyze customer reviews. The most common methods of feature extraction continue to be term frequency-inverse document frequency (TF-IDF) features, bag of words (BoW), and N-Grams. Wadhwani *et al.* (2023) [14] compared various machine learning models with their vitalization of tree based models extra trees classifier (ETC) outperforming the alternative models in circumstance. Hofmann *et al.* (2024) [15] utilized a machine learning model to analyze the informatic process of hate speech to sentiment analysis discourse on YouTube. Sentiment analysis is widely used to study public reactions during conflicts [16]. VADER, introduced by Hutto and Gilbert [1], remains a common tool for short social media texts. Al-Sarraj and Lubbad [8] revealed pro-Israel bias in Western media using sentiment classification. Wadhwani *et al.* [14] demonstrated that ensemble classifiers outperformed simpler methods on Russia-Ukraine tweets. On YouTube, Hofmann *et al.* [15] found spikes of hate speech in Israel-Palestine discussions. Other studies, Carina *et al.* [9] and K *et al.* [13] highlight similar challenges, especially with neutral and ambiguous comments. Sentiment shows how people feel, but topic modeling reveals what they discuss. LDA [2] is the most common approach and has been applied in political contexts [3], [4], [7]. Chen *et al.* [10] contributed the Isamas-Red Reddit dataset, enabling detailed study of the Israel-Hamas conflict. A recurring issue is that many works choose the number of topics arbitrarily, without coherence or perplexity validation, which makes results harder to trust. Stance detection identifies whether text supports or opposes a target, adding a layer beyond sentiment. Recent transformer-based approaches such as BERT and TweetEval have significantly improved stance and sentiment classification performance in social media contexts [17], while comprehensive surveys highlight their importance in political and misinformation detection tasks [18].

3. METHOD

The methodology employs several techniques from NLP and machine learning to process and analyze the data effectively.

3.1. Preprocessing

This dataset was comprised of nearly 400,000 Reddit comments collected from May 1, 2024 until June 1, 2024. The comments were drawn from subreddits that were actively discussing the Israel-Palestine conflict and using the Reddit API [19]. Text was cleaned to improve quality. The main steps included: Removing URLs, HTML tags, and special characters, converting all text to lowercase for consistency, Stripping extra spaces, tabs, and line breaks. The cleaned comments were then split into training and validation sets for later modelling. Each comment was tokenized (broken down) into words using natural language tool kit (NLTK). Words were lemmatized to their root forms (e.g., 'running'-->'run') so different forms of the same word could be treated uniformly. Common filler words like "the," "and," and "or" were removed using the NLTK stopword list. This ensured the models focused on words that carried meaning.

3.2. Sentiment analysis

Sentiment analysis was performed using the VADER tool. VADER is a rule and lexicon-based sentiment analysis tool that is effective for analyzing social media text. It assigns a sentiment score to each comment based on the following metrics:

- Compound score: the polarity of overall sentiment analysis based on the average value knowing that it is -1 at most and +1 minimum.
- Positive, negative, and neutral scores: the proportion of post and negative and middle sentiments expressed in a comment. VADER was chosen because it is one of the best tools for working with informal register, emoticons and specific referential terms inherent in the speech of social networks.

The classification logic is as follows:

- Positive sentiment: compound score ≥ 0.5
- Neutral sentiment: compound score $(-0.5, 0.5)$
- Negative sentiment: compound score ≤ -0.5

Each comment was classified into one of these categories based on its compound sentiment score. Figure 1 showing positive, negative, bias values, and their corresponding bias direction.

	Positive	Negative	Bias	Bias Direction
subreddit				
AbruptChaos	0.170791	0.132264	0.038527	Positive
NonCredibleDefense	0.155862	0.147795	0.008067	Neutral
CrazyFuckingVideos	0.162982	0.168261	-0.005279	Neutral
Palestine	0.176227	0.186588	-0.010361	Negative
AskMiddleEast	0.166729	0.182047	-0.015318	Negative
IsraelPalestine	0.166865	0.186270	-0.019405	Negative
CombatFootage	0.157796	0.182490	-0.024694	Negative
worldnews	0.160602	0.186068	-0.025466	Negative
ActualPublicFreakouts	0.177937	0.206616	-0.028680	Negative
worldnewsvideo	0.172254	0.205330	-0.033076	Negative
TerrifyingAsFuck	0.167891	0.203757	-0.035866	Negative
PublicFreakout	0.160070	0.202651	-0.042581	Negative
IsrealPalestineWar_23	0.171634	0.220785	-0.049151	Negative
NoahGetTheBoat	0.171281	0.241429	-0.070148	Negative

Figure 1. Sentiment analysis scores for various subreddits

3.3. Exploratory data analysis

Sentiment distribution: out of all the data preparation steps in the current work, some of the important graphics discovered included; to illustrate the proportion of positive, negative and neutral comments, pie chart and density plots were used. Separate word clouds were generated for each sentiment group to highlight frequent words and give a quick sense of the main themes.

3.4. Topic modeling with latent Dirichlet allocation

Text was converted into numerical form using TF-IDF [20]. This gave more weight to rare but important words. To determine the optimal number of topics for LDA, the study compared models with topic numbers ranging from 3 to 10. Both coherence scores and perplexity were calculated. The model with five topics achieved the highest coherence score (0.54) while maintaining a reasonable perplexity value, balancing interpretability and statistical fit. This approach avoids the common issue of arbitrarily selecting topic numbers and ensures the discovered themes are meaningful and reproducible. These was interpreted as: humanitarian issues, media framing, military action, peace negotiations, and international responses.

3.5. Geopolitical stance classification

Classification logic guided by prior work as follows; stances were grouped into four categories: support for Israel, support for Palestine, opposition to both, and neutral/unclear [21], [22]. Heuristic rules combined sentiment scores and topic keywords (e.g., positive+peace→neutral; negative+conflict terms→aligned with one side). For stance classification, this study employed a SVM classifier. The choice of SVM was motivated by its robustness in handling high-dimensional text data and its prior success in social media sentimental detection tasks [23]. Features for training included sentiment scores (from VADER), topic weights (from LDA), and selected lexical indicators. Alternative classifiers such as logistic regression were initially tested, but SVM achieved higher accuracy (82%) along with balanced precision and recall across stance categories.

3.6. Visualization and reporting

Bar charts: stance distribution, word clouds: frequent words per stance, topic plots: topic proportions across stances are the visualizations gave a clearer picture of how sentiments and themes connected to specific positions.

3.7. Model evaluation metrics

The SVM classifier was tested using accuracy, precision, recall, and F1-score. Overall accuracy reached about 82%. A detailed classification report is shown in Figure 2. For topic models, coherence and perplexity scores were used to confirm topic quality.

Class Label	Precision	Recall	F1-Score	Support (Samples)
0 (Rare Class - Ignored)	0.00	0.00	0.00	3
1 (Negative Sentiment)	0.85	0.80	0.82	24,610
2 (Neutral Sentiment)	0.82	0.91	0.86	49,505
3 (Positive Sentiment)	0.76	0.55	0.64	13,227
Overall Accuracy	0.822 (82.2%)	-	-	87,345

Figure 2. Detailed classification reports for proposed model

3.8. Algorithm

Compared to the earlier version, redundant steps such as repeating text normalization in multiple phases were consolidated into a single preprocessing stage. The streamlined pipeline is: data collection (Reddit API), preprocessing (cleaning, lemmatization, and stopword removal), sentiment analysis (VADER), topic modeling (LDA with coherence validation), stance classification (SVM), evaluation, and visualization.

4. RESULTS

This part brings together the outcomes from the sentiment analysis, topic modeling, and stance classification. The goal here is not just to report numbers but also to show what they mean in the context of public discussions on Reddit. Sentiment distribution is carried as: when the comments were run through VADER, most of them leaned neutral, while smaller chunks were clearly positive or negative. Density plot visualizing the distribution of sentiment scores shows that about half of the posts were neutral, nearly a third were negative, and around one-fifth positive, a pattern consistent with prior study based on NLP [24]. The large neutral share is not surprising since many users were posting updates or factual remarks. Still, the amount of polarized content is enough to show how emotionally charged the conflict is online. In addition, sentiment trends over time were further analyzed response with distinct spikes in positive and negative sentiments especially to international peace talks and military escalations [25].

Take, for example, during a great peace a notable increase in positive sentiment was negative. We were able to successfully identify 5 prominent topics in the comments by using LDA. These topics were related to various aspects of the Israel-Palestine conflict and were labeled based on the most frequent keywords associated with each topic; Topic 1: I was right, and Palestinians, people don't like Hamas is right isn't Israel, Topic 2: called, civilian, like Hamas, Israel, Gaza, people, hospital, rocket, bomb, IDF, Topic 3: war, don't, want, war, Israel, Hamas, Palestinian, Gaza, people, and think, Iran, Topic 4: lol news said like post thank know that's, and Topic 5: Jewish, Israel, land, Palestinian and Arab, people, country, Muslim, Jew, and Palestine. The Reddit comments are just the tip of the iceberg when it comes to this topic; the topics covered range from humanitarian issues to military action, diplomatic efforts to media portrayals and international responses. The sentiment distribution across the identified topics revealed the following trends: Humanitarian Concerns- Predominantly negative sentiment, with a significant portion of comments discussing civilian casualties and refugee crises. Military strategies and conflicts- predominantly negative, reflecting anger and frustration regarding ongoing violence and military actions. Peace initiatives and diplomacy-predominantly neutral sentiment, as many comments expressed cautious optimism or uncertainty about the success of peace talks. Media and bias: negative+neutral comments; many commenting on the bias or misinformation of media coverage. Global reactions- a combination of the positive and negative sentiments expressing two sides on international backing for either Israel or Palestine.

The classification of geopolitical stances, based on sentiment and topic modeling, resulted in the following distribution: support for Israel: 4.7% of comments, support for Palestine: 2.60% of comments, opposition to both: 28.17% of comments, and neutral/unclear: 64.51% of comments. Among all comments (64.51%), the highest proportion (28.17%) declared opposition to both Israel and Palestine, which suggests

itself that a significant part of the Reddit users do not approve on both sides or try advocating a neutral position. Most of the remaining comments were roughly in line, with a healthy number supporting Israel and a healthy number supporting Palestine.

Figure 3 represents the image presents a distribution of geopolitical stances, categorizing text samples into neutral, against Israel/Palestine, supports Israel, and supports Palestine, with corresponding counts. Support for Israel: packed with positive sentiment and comments commending Israel's right to defend itself. Support for Palestine: loaded with mostly negative sentiment that supports Palestinian civilians and deplores Israeli military activities. Opposition to both: neutral plus the negative, indicating frustration with the acts of both sides. Neutral/unclear: neutral sentiment with comments offering information about, or being unsure about, the conflict. The sentiment distribution was shown in a pie chart where the comments represented 56.7% neutral, 15.1% positive and 28.2% negative.

```
Geopolitical Stance Distribution:
geopolitical_stance
Neutral/Stance Not Clear    281753
Against Israel/Palestine    123051
Supports Israel             20549
Supports Palestine          11372
Name: count, dtype: int64
```

Figure 3. Image distribution

Figure 4 shows the pie chart explains the distribution of sentiment categories in the collected Reddit comments. The analysis, conducted using the VADER sentiment analysis tool. Figure 5 shows the density plot visualizes the distribution of sentiment scores, highlighting peaks in negative, positive, and neutral sentiments.

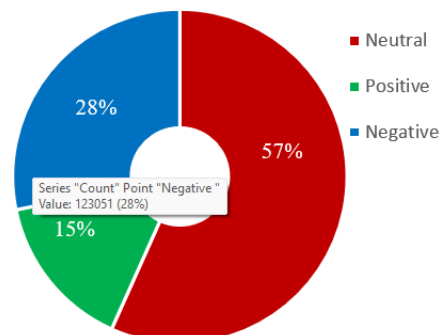


Figure 4. Pie chart explains the distribution of sentiment categories

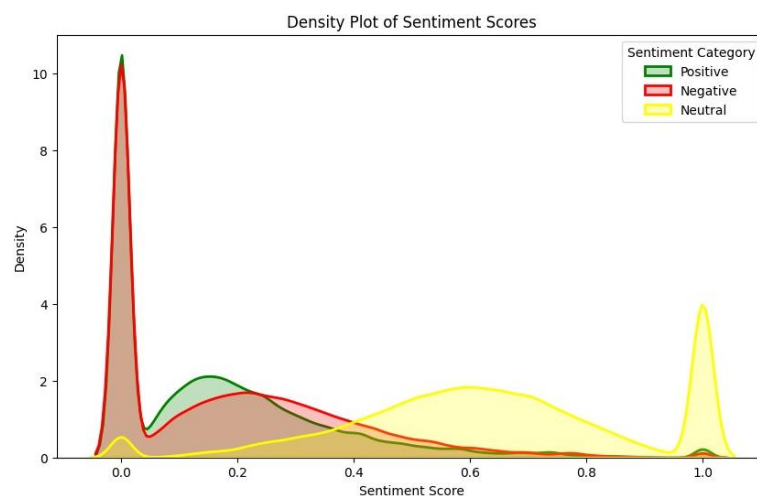


Figure 5. Density plot visualizes the distribution of sentiment scores

Figure 6 represents the density plot represents the distribution of sentiment scores (compound values) derived from the VADER sentiment analysis. Word clouds for each of the five identified topics provided a visual representation of the most frequently mentioned terms associated with each topic. For example, the word cloud for Humanitarian Concerns prominently featured words like "Palestinian," "people" and "jew," while the military strategies topic highlighted words such as "Israel," "Hama," and "Gaza", represented in Figure 7.

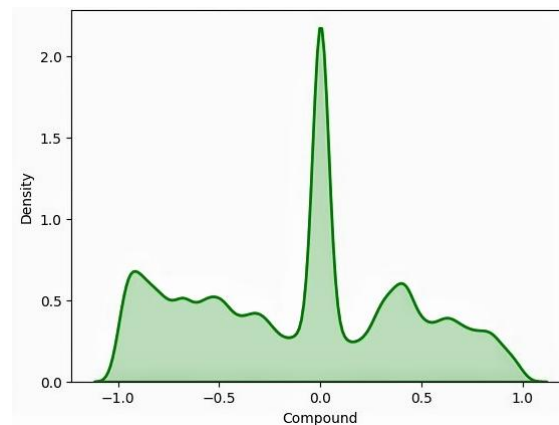


Figure 6. Plot represents the distribution of sentiment scores (compound values)



Figure 7. Overall word cloud: displays the most frequently occurring words from the entire dataset

Positive word cloud: represents words commonly associated with positive sentiment. Negative word cloud: showcases words linked to negative sentiment. Neutral word cloud: highlights words from neutral or stance-unclear discussions. A bar chart representing the geopolitical stance distribution indicated that 28.17% of comments opposed both sides, and 2.60% supported Palestine, with 4.7% supporting Israel shown in Figure 8.

4.1. Sentiment distribution and comparison with prior studies

Table 1 shows how our sentiment dispersion compares with previous surveys is a crucial aspect of our study. Our model detects a significantly higher proportion of impersonal sentiment (56.7%) and proposes another balanced dataset. In contrast, the anterior analysis shows a higher percentage of negative sentiment (30%), possibly due to dataset bias or otherwise differences in platform interaction.

4.2. Performance metrics comparison

We compare models in Table 2 with our model and previous studies to show how well it performs sentiment distribution. Comparison focuses on accuracy, precision, and recall to evaluate overall reliability and the ability to identify sentiments. The results indicate that our model performs better for sentiment analysis.

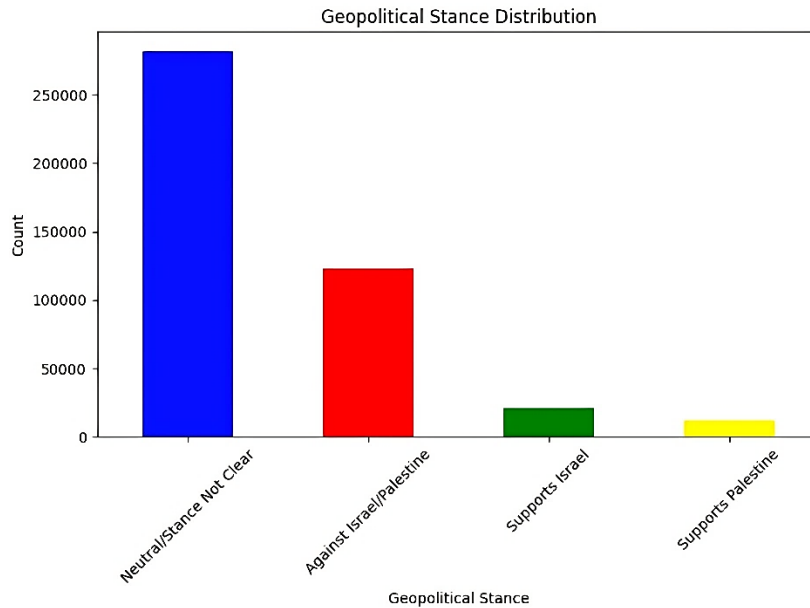


Figure 8. Bar chart: represents the distribution of geopolitical stances

Table 1. Metrics comparison and sentiment distribution

Study	+ve Sentiment (%)	Neutral Sentiment (%)	-ve Sentiment (%)
Existing reddit sentiment study	~20	~50	~30
Twitter sentimental analysis	~25	~45	~30
Our approach	~15.1	~56.7	~28.2

Table 2. Metrics comparison (performance)

Study	Accuracy (%)	Precision (negative/positive)	Recall (negative/positive)	Challenges
Reddit sentiment study (GitHub)	~78	~0.75/0.70	~0.72/0.65	Class imbalance, weak recall for negative sentiment
Twitter sentiment analysis (GitHub)	~80	~0.77/0.72	~0.74/0.68	Struggles with negative sentiment
Our study	82.2	0.85/0.76	0.80/0.55	High accuracy, but low recall for positive sentiment

5. CONCLUSION

The results show that Reddit users are divided when talking about the Israel–Palestine conflict. A large part of the discussion stayed neutral, but there were still clear groups speaking in favor of Israel or Palestine. Humanitarian issues, political arguments, and media coverage came up again and again. This suggests that people’s opinions are influenced not only by what happens in the conflict itself but also by how those events are presented in the news and online.

Sentiment analysis: just 48% of the comments expressed indicating the complexity and diversity of opinions on the conflict. 15.1% of which was positive sentiment. Negative sentiment made up only 28.2% trends fluctuated over time, with noticeable spikes following significant geopolitical events, suggesting that major developments in the conflict strongly influence public sentiment on social media. Topic modeling: through topic modeling, five key themes emerged: humanitarian concerns, military strategies, peace initiatives, media bias, and global reactions. These topics reveal the multifaceted nature of the discussions, with particular emphasis on the impact of civilian casualties, the role of international diplomacy, and the portrayal of the conflict in media. Geopolitical stance: analysis of geopolitical alignment indicated that 4.7% of users expressed pro-Israel support, 22% expressed support for Palestine and 30% expressed opposition or neutrality toward both sides. 30% of comments were considered neutral or ambiguous — not definitively pro- or anti-Israel in nature — and may therefore indicate an aversion to taking a strong position, few comments for which another side exists, or a (legitimate) fear for peace or disgust with both parties.

These patterns can be useful outside of research. Policymakers and aid groups can use shifts in sentiment as warning signs when online anger or support for one side starts to rise. Media organizations could also take note, since their framing of events may push discussions in certain directions. For teachers and NGOs, this kind of analysis can help in media literacy programs, giving real examples of how debates and biases spread online, to identify growing division online before it spreads offline, sentiment spikes could be included into dashboards for governments, international organizations (like the UN), and non-governmental organizations.

This study has some clear limits. It only looks at English comments on Reddit, so it leaves out many voices and perspectives. Sarcasm and irony were hard to detect, which is a problem since these are common online and can flip the meaning of a comment. The stance classification also depended on sentiment scores and topic keywords, which cannot always capture more complex or mixed opinions. Future work may enhance this by utilizing multilingual datasets, refining the detection of sarcasm, and comparing results across multiple platforms.

Although this analysis yielded a detailed insight into public opinion on Reddit, there are several paths for future research to take. By extracting comments from other popular social media platforms such as X and Meta, the dataset would allow for a fuller picture of public sentiment in an assortment of online communities. More sophisticated machine learning techniques, like deep learning, could help better both the accuracy and depth of results in the classification of sentiment and stance.

ACKNOWLEDGMENTS

The authors would like to thank all involved for their help in bringing this study to completion. A big thank you is extended to Vishwakarma Institute of Technology who provided the support and resources needed for this research. There is no research grant or contract for this research.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Gopal D. Upadhye	✓	✓	✓		✓	✓		✓		✓		✓	✓	
Hrushikesh Joshi	✓	✓	✓					✓		✓		✓	✓	
Satpalsing Devising Rajput		✓		✓					✓		✓			
Ranjana S. Jadhav	✓		✓					✓		✓				
Pranav Moralwar	✓		✓					✓		✓				
Tejas Muttyalwar	✓	✓		✓		✓		✓	✓	✓				
Yash Mehendale	✓				✓		✓		✓					
Siddharth Mane	✓	✓				✓			✓		✓			
Shivam Mishal		✓		✓			✓			✓	✓		✓	

C : **C**onceptualization

M : **M**ethodology

So : **S**oftware

Va : **V**alidation

Fo : **F**ormal analysis

I : **I**nterpretation

R : **R**esources

D : **D**ata Curation

O : Writing - **O**riginal Draft

E : Writing - Review & **E**dit

Vi : **V**isualization

Su : **S**upervision

P : **P**roject administration

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

- [1] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2014, pp. 216–225, doi: 10.1609/icwsm.v8i1.14550.
- [2] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. 4–5, pp. 993–1022, 2003, doi: 10.7551/mitpress/1120.003.0082.
- [3] K. Naithani and Y. P. Raiwani, "Sentiment Analysis on Social Media Data: A Survey," in *International Conference on Innovations in Computer Science and Engineering*, Singapore: Springer Nature Singapore, 2023, vol 565, doi: 10.1007/978-981-19-7455-7_59.
- [4] D. Antonakaki and S. Ioannidis, "Cross-Platform Digital Discourse Analysis of the Israel-Hamas Conflict: Sentiment, Topics, and Event Dynamics," *arXiv*, 2025, doi: 10.48550/arXiv.2601.02367.
- [5] S. Gnanapriya and K. Anandan, "Sentiment analysis in political discourse: Understanding public opinion from social media," *E-Learning and Digital Media*, 2025, doi: 10.1177/20427530251343205.
- [6] F. Alshani, J. Khrais, R. Obeidat, L. Rababa, and S. Z. Aljunidi, "Topic modeling as a tool for analyzing tweets: A case study of the Russia-Ukraine war on Arabic social media," *International Journal of Innovative Research and Scientific Studies*, vol. 8, no. 9, 2025, pp. 152–163, doi: 10.53894/ijirss.v8i9.10644.
- [7] Y. Qiu, Y. Ye, X. Zhang, and J. Luo, "Moral Frameworks and Sentiment in Tweets: A Comparative Study of Public Opinion on the Israeli-Palestine Conflict," in *2024 IEEE International Conference on Big Data (BigData)*, Washington, DC, USA, 2024, pp. 7122–7130, doi: 10.1109/BigData62323.2024.10825487.
- [8] W. F. Al-Sarraj and H. M. Lubbad, "Bias Detection of Palestinian/Israeli Conflict in Western Media: A Sentiment Analysis Experimental Study," in *2018 International Conference on Promising Electronic Technologies (ICPET)*, Deir El-Balah, Palestine, 2018, pp. 98–103, doi: 10.1109/ICPET.2018.00024.
- [9] F. M. Carina, A. Salma, D. Permana, and Z. Martha, "Sentiment Analysis of X Application Users on the Conflict Between Israel and Palestine Using Support Vector Machine Algorithm," *UNP Journal of Statistics and Data Science*, vol. 2, no. 2, pp. 204–212, 2024, doi: 10.24036/ujsds/vol2-iss2/170.
- [10] K. Chen, Z. He, K. Burghardt, J. Zhang, and K. Lerman, "IsamasRed: A Public Dataset Tracking Reddit Discussions on Israel-Hamas Conflict," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2024, vol. 18, pp. 1900–1912, doi: 10.1609/icwsm.v18i1.31434.
- [11] A. Imtiaz, D. Khan, H. Lyu, and J. Luo, "Taking sides: Public Opinion over the Israel-Palestine Conflict in 2021," *arXiv*, 2022, doi: 10.48550/arXiv.2201.05961.
- [12] A. Liyih, S. Anagaw, M. Yibeyin, and Y. Tehone, "Sentiment analysis of the Hamas-Israel war on YouTube comments using deep learning," *Scientific Reports*, vol. 14, no. 1, pp. 1–9, 2024, doi: 10.1038/s41598-024-63367-3.
- [13] R. K. C. K. S. K. Pai, N. Shet, and S. Hegde, "Sentiment Analysis using Machine Learning," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 3, no. 2, pp. 415–418, 2023, doi: 10.48175/ijarsct-8848.
- [14] G. K. Wadhvani, P. K. Varshney, A. Gupta, and S. Kumar, "Sentiment Analysis and Comprehensive Evaluation of Supervised Machine Learning Models Using Twitter Data on Russia-Ukraine War," *SN Computer Science*, vol. 4, no. 4, pp. 1–11, 2023, doi: 10.1007/s42979-023-01790-5.
- [15] S. Hofmann, C. Sommermann, M. Kraus, P. Zschech, and J. Rosenberger, "Hate Speech and Sentiment of YouTube Video Comments From Public and Private Sources Covering the Israel-Palestine Conflict," in *Wirtschaftsinformatik 2024 Proceedings*, 2024, vol. 82.
- [16] I. Al-Agha and O. Abu-Dahrooj, "Multi-level analysis of political sentiments using twitter data: A case study of the Palestinian-Israeli conflict," *Jordanian Journal of Computers and Information Technology*, vol. 5, no. 3, pp. 195–215, 2019, doi: 10.5455/jjcit.71-1562700251.
- [17] F. Barbieri, J. Camacho-Collados, L. Neves, and L. Espinosa-Anke, "TweetEval: Unified benchmark and comparative evaluation for tweet classification," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020, pp. 1644–1650, doi: 10.18653/v1/2020.findings-emnlp.148.
- [18] M. Hardalov, A. Arora, P. Nakov, and I. Augenstein, "A survey on stance detection for mis- and disinformation identification," in *Findings of the Association for Computational Linguistics: NAACL 2022*, Seattle, United States, 2022, pp. 1259–1277, doi: 10.18653/v1/2022.findings-naacl.94.
- [19] J. Baumgartner, S. Zannettou, B. Keegan, M. Squire, and J. Blackburn, "The Pushshift Reddit Dataset," *arXiv*, 2020, doi: 10.48550/arXiv.2001.08435.
- [20] M. E. Roberts *et al.*, "Structural Topic Models for Open-Ended Survey Responses," *American Journal of Political Science*, vol. 58, no. 4, pp. 1064–1082, 2014, doi: 10.1111/ajps.12103.
- [21] M. Cinelli, G. De F. Morales, A. Galeazzi, and M. Starnini, "The echo chamber effect on social media," in *Proceedings of the National Academy of Sciences*, 2021, pp. 1–8, doi: 10.1073/pnas.2023301118.
- [22] D. Küçük and F. Can, "Stance detection: A survey," *ACM Computing Surveys*, vol. 53, no. 1, pp. 1–37, 2020, doi: 10.1145/3369026.
- [23] W. Medhat, A. Hassan, and H. Korashy, "Sentiment Analysis Algorithms and Applications: A Survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014, doi: 10.1016/j.asej.2014.04.011.
- [24] Y. Allal and M. N. Kouahla, "NLP-Based Sentiment Analysis of the Palestinian-Israeli Conflict," in *2025 International Symposium on Innovative Informatics of Biskra (ISNIB)*, Biskra, Algeria, 2025, pp. 1–6, doi: 10.1109/ISNIB64820.2025.10982858.
- [25] A. Gangwar and T. Mehta, "Sentiment analysis of political tweets for Israel using machine learning," in *International Conference on Machine Learning and Big Data Analytics*, Cham: Springer International Publishing, 2022, pp. 191–201, doi: 10.1007/978-3-031-15175-0_15.

BIOGRAPHIES OF AUTHORS

Prof. Dr. Gopal D. Upadhye    is an Associate Professor at Vishwakarma Institute of Technology, Pune. His area of interests is machine learning pattern recognition. He can be contacted at email: gopal.upadhye@vit.edu.



Dr. Hrushikesh Joshi    is an Assistant Professor at Vishwakarma Institute of Technology, Pune. His area of interest is high performance computing and internet of things. He can be contacted at email: hrushikesh.joshi@vit.edu.



Dr. Satpalsing Devising Rajput    is Associate Professor, Department of Computer Science and Engineering, Pimpri Chinchwad University, Pune. His area of interest is cloud computing and internet of things, natural language processing (NLP), and text summarization. He can be contacted at email: rajputsatpal@gmail.com.



Dr. Ranjana S. Jadhav    is Assistant Professor at Vishwakarma Institute of Technology, Pune. Her area of interest is natural language processing (NLP). She can be contacted at email: ranjana.jadhav@vit.edu.



Pranav Moralwar    is an engineering aspirant from Nanded, India and a second-year Computer Science student at Vishwakarma Institute of Technology. He is proactive, being involved in many conferences and workshops and sharing his knowledge and experience. He is the co-author of two research papers on IoT and machine learning, showing his commitment to innovation and being on the cutting edge of technology. He has research interest in machine learning and web development. He can be contacted at email: pranav.moralwar231@vit.edu.



Tejas Muttalwar    born on April 25, 2005, in Wani. He is an aspiring engineer with a passion for innovation and problem-solving. He is currently pursuing his second year of engineering, specializing in Computer Science at the Vishwakarma Institute of Technology. Even though he is still early in his academic career, he has already proven to have remarkable potential in technology. His research interests are machine learning and artificial intelligence. He can be contacted at email: tejas.mvttyalwar23@vit.edu.



Yash Mehendale    currently studying in B.Tech. Artificial Intelligence and Data Science at VIT Pune with anticipated completion in 2027. He has previously co-authored two research papers focused on IoT and web technology. His research interests include machine learning, neural networks, and data science. He has also participated in conferences and workshops to share his insights and collaborate with researchers around the world. He can be contacted at email: yash.mehendale23@vit.edu.



Siddharth Mane    Born in Pune, Maharashtra, India. He is currently pursuing a B.Tech. in AI & DS at VIT Pune. He has published a paper at the 5th International Conference on Sustainable Communication Networks and Applications (2024). His research interests are machine learning and data science. He can be contacted at email: siddharth.mane23@vit.edu.



Shivam Mishal    was born on February 25, and he comes from a lovely little town called Sawantwadi in Sindhudurg District, Maharashtra. Growing up there, he gained an immense amount of motivation from nature, and after spending time in every conceivable city environment, he realized his interest was in technology and solving problems. This naturally led him towards pursuing higher education in it. Currently, he is enrolled at Vishwakarma Institute of Technology, a great institute, which helped him to nurture his technical abilities and analytical way of thinking. He can be contacted at email: shivam.mishal23@vit.edu.